Obtaining Reliable Measures at the Physician-level: Some Important Methodological Issues

Dana Gelb Safran, ScD
The Health Institute
Institute for Clinical Research and Health Policy Studies
Tufts-New England Medical Center

Presented at:
The Ninth National CAHPS User Group Meeting
Baltimore, MD

2 December 2004

Focusing on Physicians

- Survey-based measurement of patients' experiences with individual physicians is not new.
- What's new: Efforts to standardize and potential for public reporting.
- IOM report *Crossing the Quality Chasm* gave "patient-centered care" a front row seat.
- Methods and metrics have been honed through 15 years of research and through several recent large-scale demonstration projects
- But putting these measures to use raises many questions about feasibility and value.

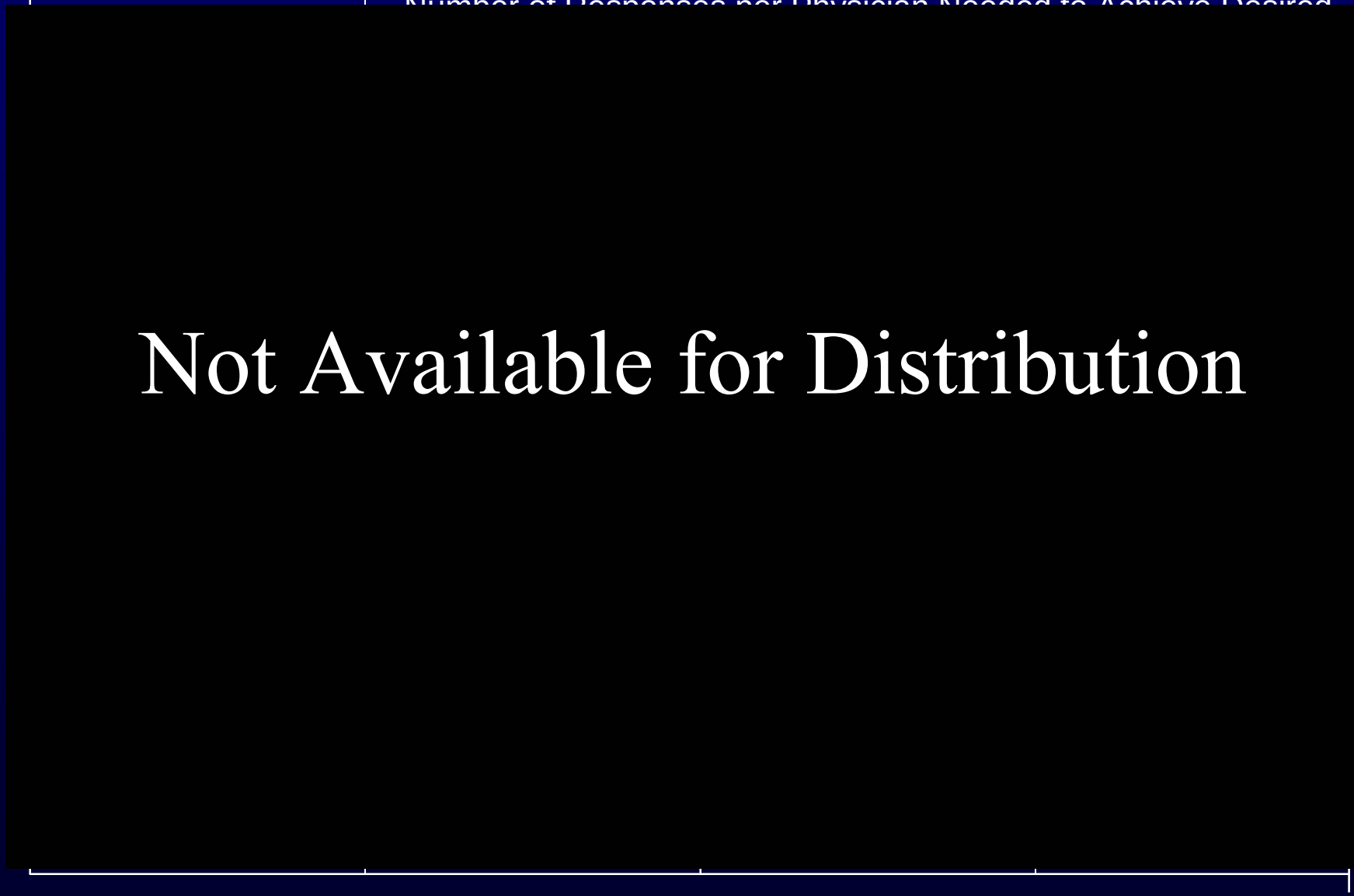
Some "1st Generation" Questions of Moving MD-Level Measurement into Practice

- What sample size is needed for highly reliable estimate of patients' experiences with a physician?
- What is the risk of misclassification under varying reporting frameworks?
- Is there enough performance variability to justify measurement?
- How much of the measurement variance is accounted for by physicians as opposed to other elements of the system (practice site, network organization, plan)?

Some "2nd Generation" Questions of Moving MD-Level Measurement into Practice

- Can the data be obtained with methods that are less costly (and more flexible) that "mail"?
- How do other modes (particularly internet and interactive voice response telephone [IVR]) affect response rates and data quality.
- What does it take to improve performance on these measures?

Sample Size Requirements for Varying Physician-Level Reliability Thresholds



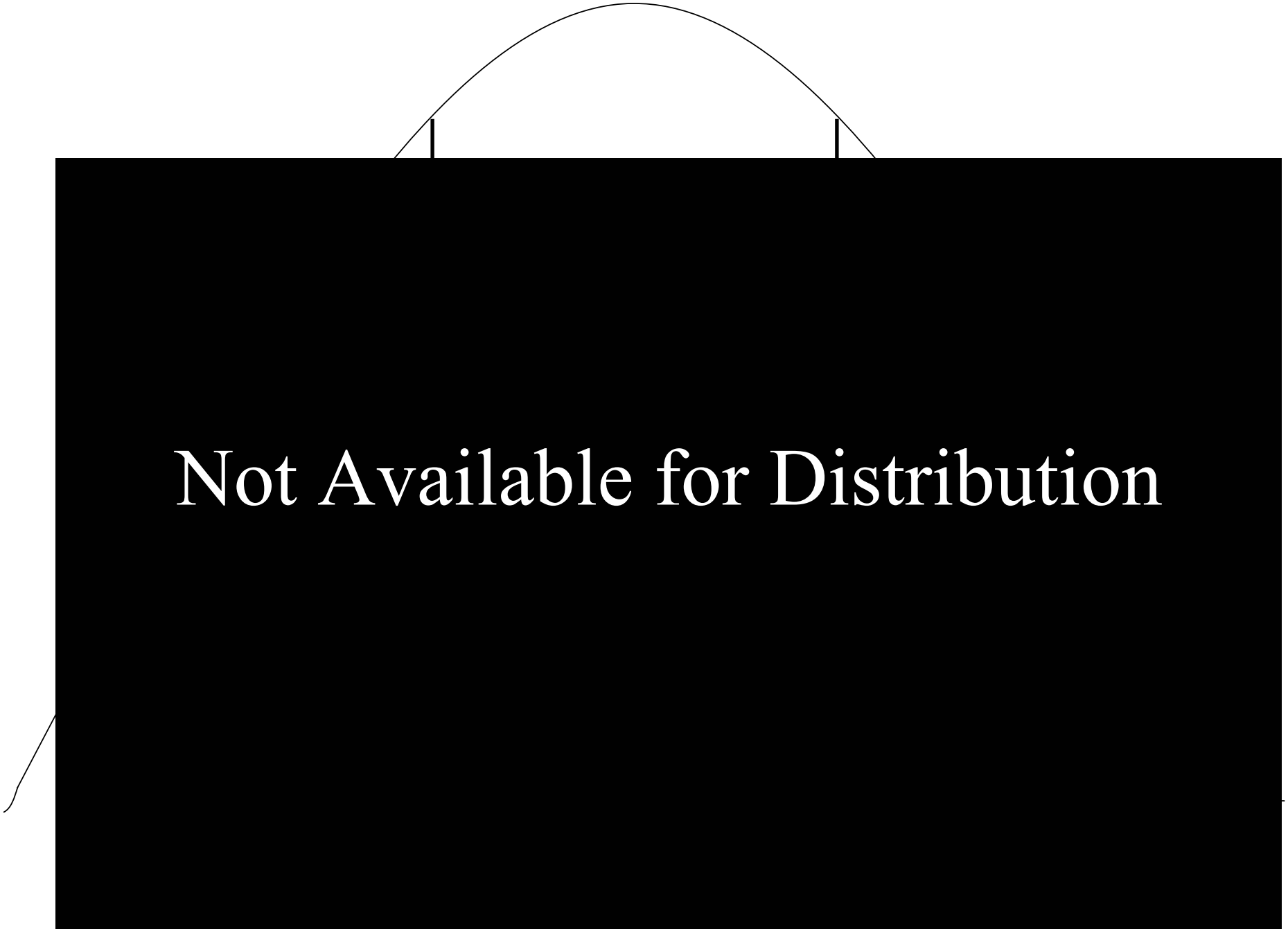
What is the Risk of Misclassification?

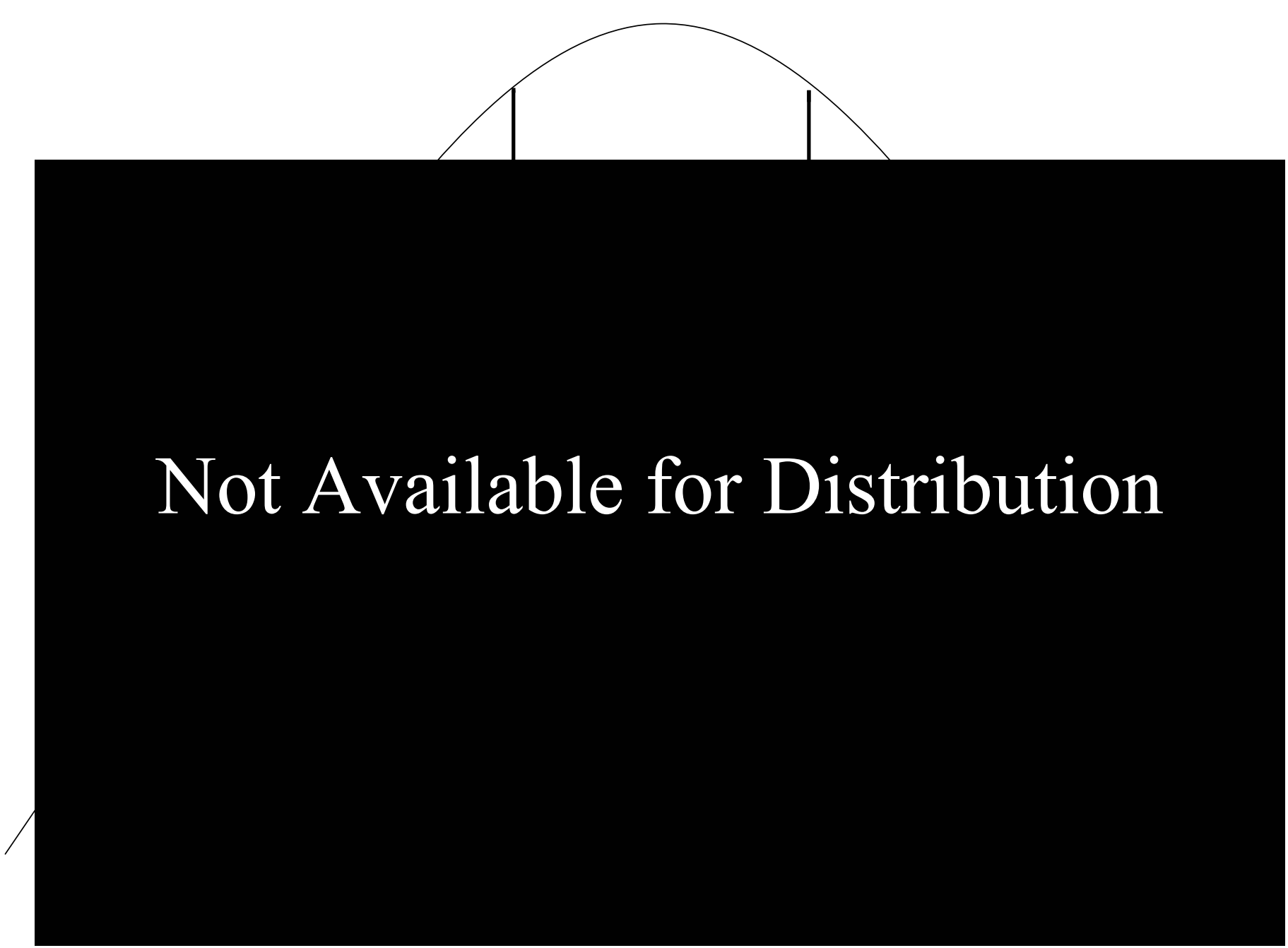
- Not simply 1- α_{MD}
- Depends on:
  - Measurement reliability (α_{MD})
  - Proximity of score to the cutpoint
  - Number of cutpoints in the reporting framework

Risk of Misclassification at Varying Distances from the Benchmark and Varying in Measurement Reliability (α_{MD})

Not Available for Distribution

Certainty and Uncertainty in Classification


Not Available for Distribution


Not Available for Distribution

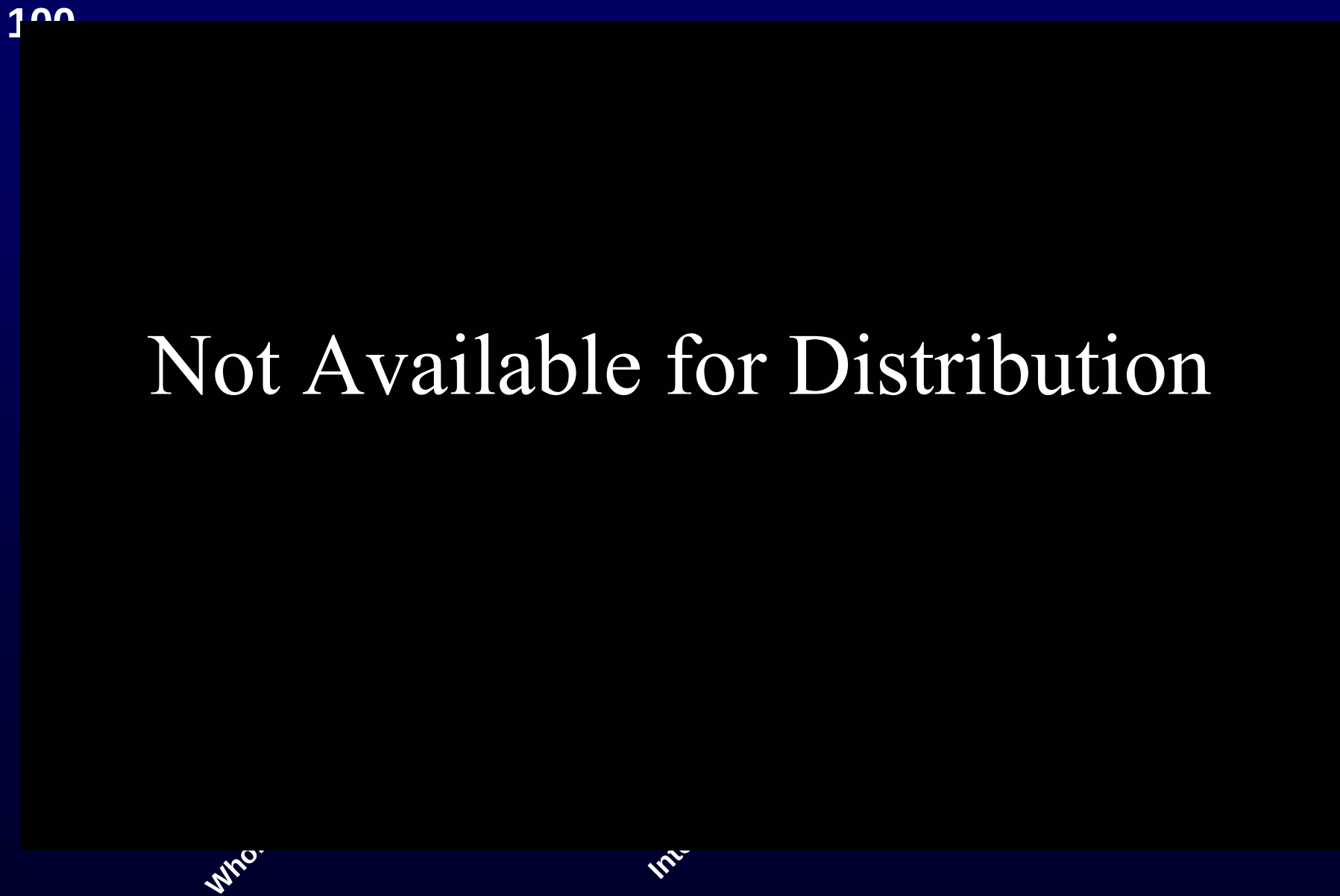
Variability Among Physicians (Communication)

Not Available for Distribution

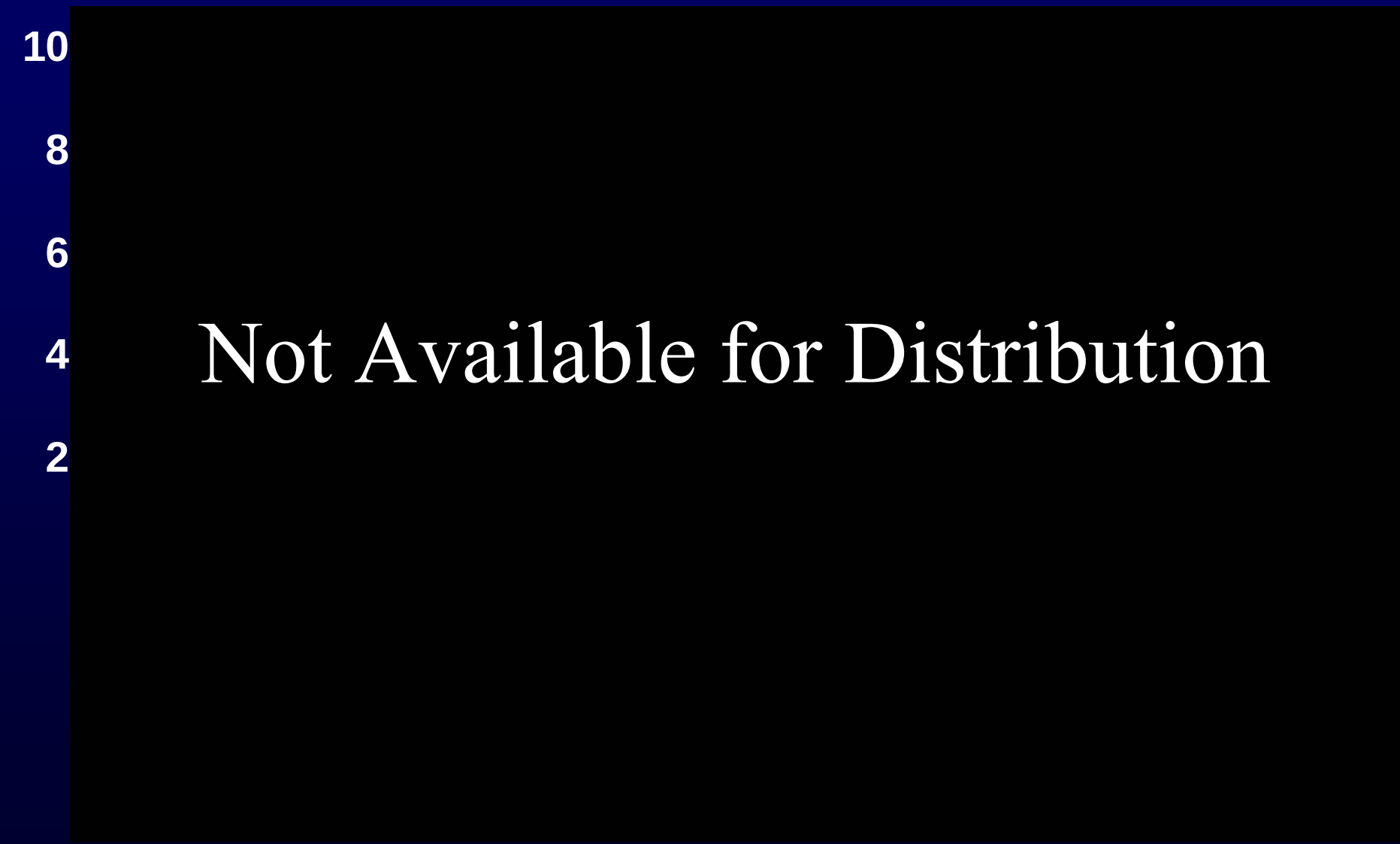
Variability Across Practice Sites (Communication)

Variability Among Physicians within Sites (Communication)

Allocation of Explainable Variance: Doctor-Patient Interactions



Allocation of Explainable Variance: Organizational/Structural Features of Care



Mode Trial Results: Response Rates

	Mail	Web		IVR	
		Web Alone	Web + Mail X-over	IVR Alone	IVR + Mail X-over
Total	**49.6**	**17.2**	**45.6**	**30.8**	**49.3**
Group 1	43.0	15.3	37.6	22.5	40.1
Group 2	49.3	15.7	44.5	30.9	48.5
Group 3	54.5	30.5	51.2	40.5	56.3
Group 4	58.8	24.9	56.2	38.3	61.5
Group 5	48.8	9.4	45.0	29.9	52.1

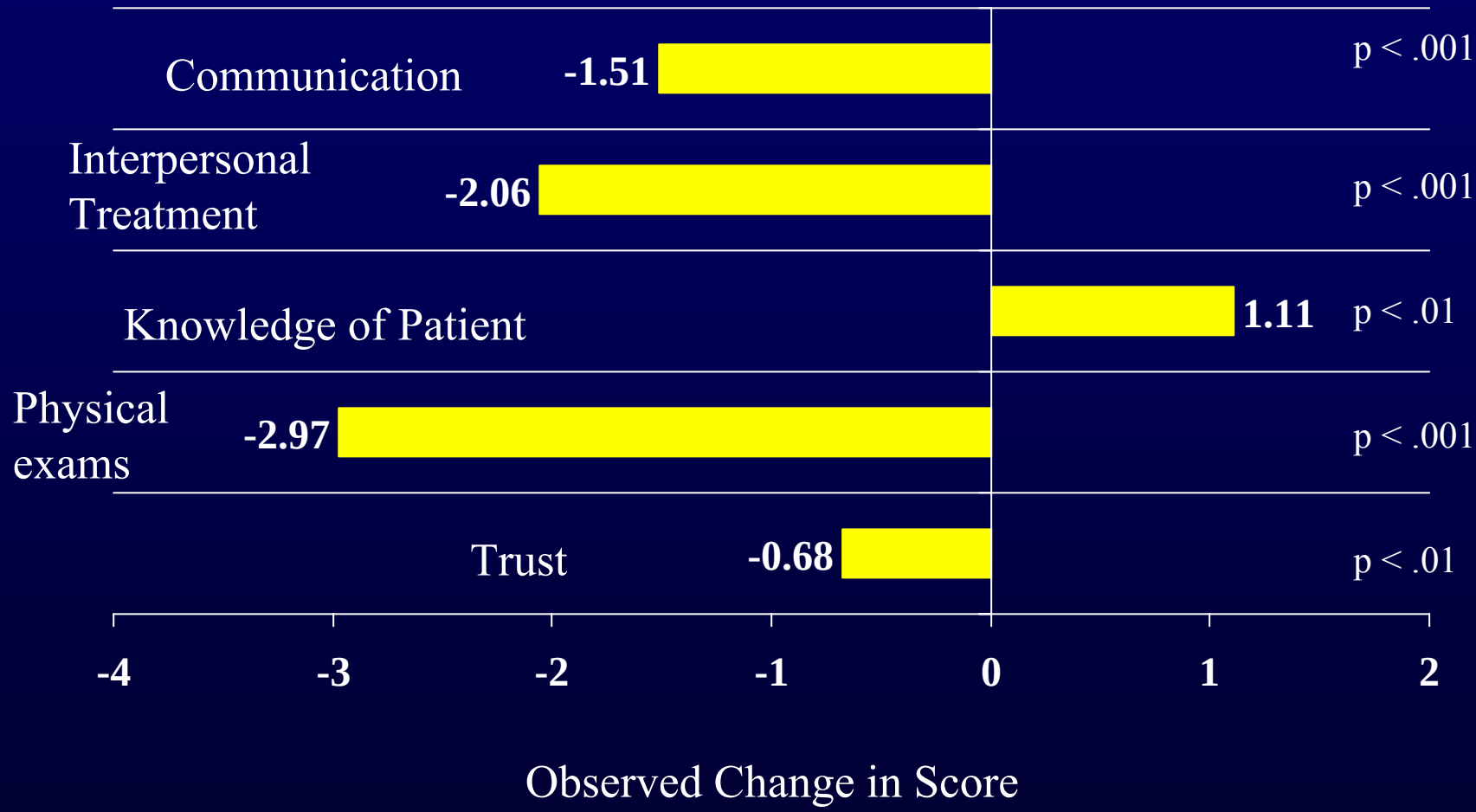
NOTE: The denominators used in calculating the response rates do not exclude ineligibles (e.g., death, bad address).

Mode Trial Results: Comparison of 3 Modes, Unadjusted and Adjusted

	Unadjusted			Adjusted		
	Mail	Web_Only	IVR_Only	Mail	Web_Only	IVR_Only
	n = 2362	n = 1477	n = 1960	n = 2362	n = 1477	n = 1960
Quality of MD-Pt Interaction	85.0	84.8	**82.5**	73.4	73.2	**70.8**
Coordination	76.6	77.1	**72.8**	63.3	64.1	**59.2**
Access/ Continuity	77.9	78.9	**70.2**	64.7	65.8	**57.1**
Recommend MD	89.4	88.7	90.6	79.4	78.6	80.2

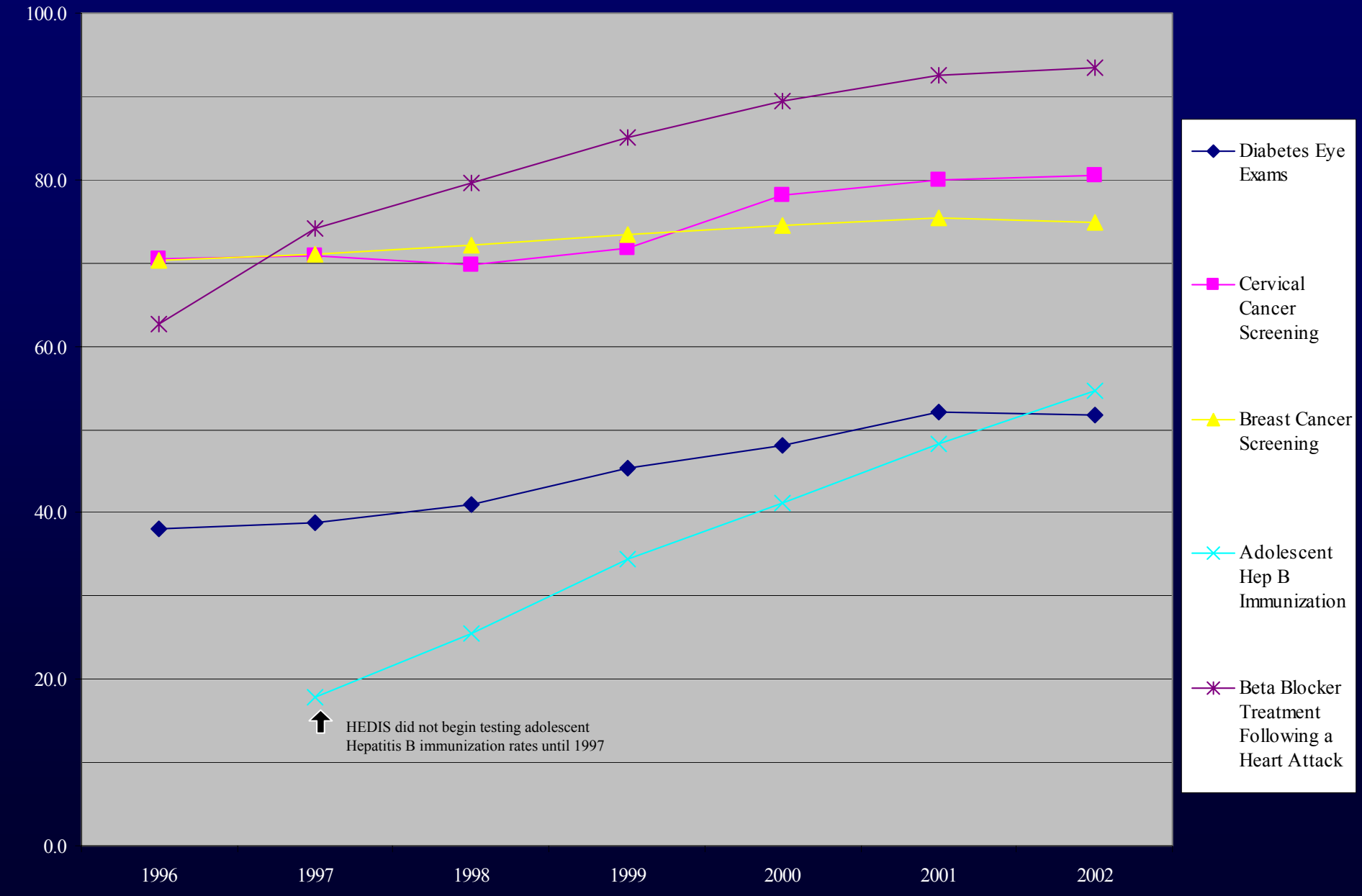
Note: **Results in** bold **denote those that are statistically significantly different from Mail ($p \leq .05$). Results that are underlined denote those for which there are statistically significantly differences between Web+Mail and IVR+Mail.**

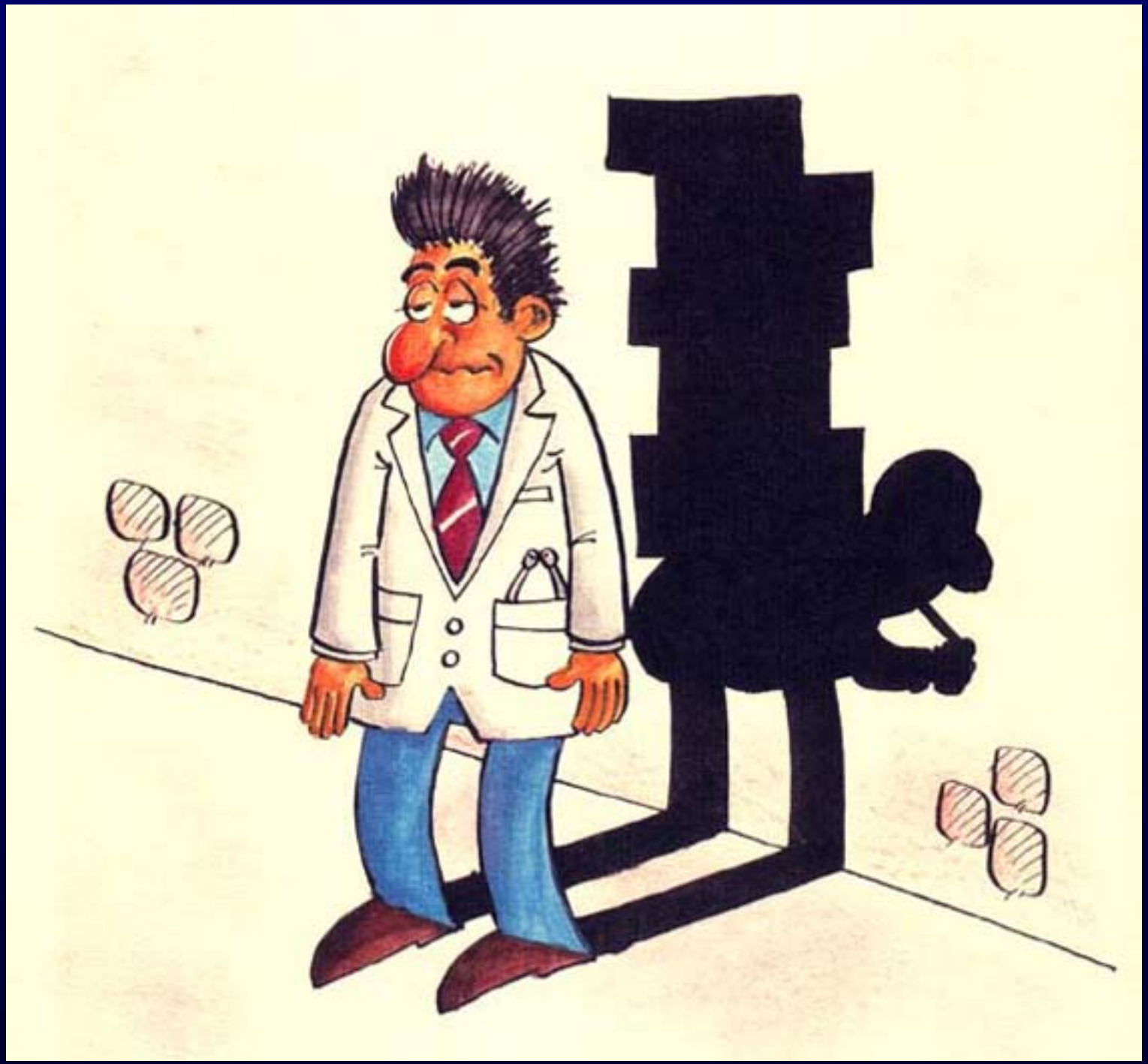
Primary Care Relationship Quality & Interactions, 1996-1999



Source: Murphy et al. JFP 2001.

Changing Rates of Preventive Care Processes, 1996-2001





Doctor and the Doll by Norman Rockwell